REMARKS

I. INTRODUCTION

In response to the Office Action dated June 16, 2006, claims 1-9, 15-23 and 29-37 have been amended. Claims 1-42 remain in the application. Re-examination and re-consideration of the application, as amended, is requested.

II. DOUBLE PATENTING REJECTIONS

In paragraphs (4)-(5) of the Office Action, claims 1-42 were rejected on the grounds of nonstatutory obviousness-type double patenting as being unpatentable over claims 1-45 of U.S. Patent No. 6,954,758.

Applicant's attorney notes the provisional nature of these rejections, and will substantively address these rejections upon an indication of otherwise allowable claims.

III. PRIOR ART REJECTIONS

A. The Office Action Rejections

In paragraphs (2)-(3) of the Office Action, claims 1-42 were rejected under 35 U.S.C. §102(b) as being anticipated by Pham, U.S. Patent No. 5,970,482 (Pham).

Applicant's attorney respectfully traverses these rejections.

B. Applicant's Independent Claims

Applicant's independent claims 1, 15 and 29 are directed to a method, apparatus and article of manufacture for using predictive models within a computer-implemented business analysis environment. Claim 1 is representative, and comprises the steps of:

(a) applying a derived measure against a segment, wherein the derived measure comprises a predictive model previously-built by a model-building mechanism in a data mining system, wherein the derived measure is invoked within an application template that is a sequence of segments, filters, measures and functions linked together in a workflow; and

(b) generating output for the segment from the predictive model in the form of measure values.

C. The Pham Reference

Pham describes a neuroagent approach that is used in an automated and unified data mining system to provide an explicitly predictive knowledge model. The neuroagent is a neural multi-agent approach based on macro-connectionism and comprises a double integration at the association and symbolic level as well as the knowledge model level. This data mining system permits discovery, evaluation and prediction of the correlative factors of data, i.e., the conjunctions, as corresponding to neuroexpressions (a semantic connection of neuroagents) connected to an output neuroagent which corresponds to the data output, the connection weights yielding the relative significance of these factors to the given output. The system takes data sets called Domains, establishes candidate dimensions or Parameters, categorizes Parameters into discrete bins, and trains a neuroagent network composed of neuroagents allocated for each bin and each output based on a discovery data set, called a Discovery Domain, and by building up the various minimal and contextual neuroexpressions, and setting the appropriate connection weights, the results may therefore be compared with an optional evaluation data set, called an Evaluation Domain to establish the accuracy of the knowledge model, and thereafter applied with some degree of confidence to a prediction set or Prediction Domain. The ranking in importance of the composite Parameters may be calculated as well as the discrimination between the various outputs, which permits the relevant factors of interest to a decision maker to come into focus.

D. The Applicant's Independent Claims Are Patentable Over The Reference

Applicant's independent claims 1, 15 and 29 are patentable over the reference because they recite a novel and nonobvious combination of steps and elements. Specifically, Applicant's independent claims 1, 15 and 29 have been amended to recite that "the derived measure is invoked within an application template that is a sequence of segments, filters, measures and functions linked together in a workflow." These limitations were originally found in dependent claims 2, 16 and 30.

The Office Action asserted that these limitations could be found in Pham at various locations:

> [Claim 16] wherein the derived measure is invoked within an application template, the application template comprises a sequence of elements linked together in a workflow, and the elements are-selected from a group -- - comprising a segment, a filter, a measure and a function (col. 13, lines 1-12, 40-55; col. 14, lines 58-67 -- The Knowledge Model engine generates a knowledge model based on data mining techniques; col. 19, lines 4-52; col. 20, lines 8-67 -- The knowledge model is applied

to categories or bins of data; col. 9, line 41 through col. 12, line 54 -- Neuroagents, programmed with various neuroexpressions, may be used during the learning and training phases. In order to accomplish this functionality, a workflow linking the elements of segment, filter, measure, and function is carried out);

At the indicated locations, Pham describes the following:

<u>Col. 13, lines 1-12</u>

An automatic pattern discovery process is needed. Pattern discovery is the capability to establish relationships between different pieces of information (features) to define a pattern which in turn represents a context capable of identifying a situation or defining it. The adjective "automatic" excludes the classic analysis tools (fast data access tools, DSS, EIS, OLAP, etc.) for the development of "intelligent" Data Mining systems, as shall be discussed below, because with those tools one must know in advance exactly what one must search. These tools thus act as a "verification" mode. This constitutes the main limitation of these kind of tools in exploring complex data sets.

<u>Col. 13, lines 40-55</u>

Data mining tools extend normal thinking logic by analyzing complex data automatically in order to build a Knowledge Model useful for both understanding and prediction. The present invention surpasses other Data Mining tools. None of the other technologies used in current products (Neural Networks, rule induction, Case-Based Reasoning and statistical analysis) can provide the combination of prediction, explanation, performance and ease of use necessary for widespread information discovery. The present invention is a unified approach towards Data Mining. Where other prior art systems may combine some of these other technologies in a hybrid system, one component in the hybrid providing the prediction, another component the explanation, the knowledge model of the present invention provides, in a unified manner, both prediction and explanation.

<u>Col. 14, lines 58-67</u>

These and other objects of the present invention are accomplished by using a computer-implemented data mining system comprised of a Study Manager, Discovery Manager, Evaluation Manager, Prediction Manager and Knowledge Model engine. The Study Manager allows a Discovery Domain, an Evaluation Domain and a Prediction Domain to be selected from one or more Data Sources, each Data Source including one or more data records, and each data record having one or more Parameters. The Knowledge Model engine, when presented with the Discovery Domain, constructs an explicitly predictive Knowledge Model therefrom and returns a discovery results set, and when presented with either of the Evaluation or Prediction Domains, applies the Knowledge Model thereto and returns an evaluation or prediction results set, respectively. The Discovery Manager takes the discovery results set from the Knowledge Model engine and calculates the relative significance of the Parameters under the Knowledge Model. The Evaluation Manager takes the evaluation results set from the Knowledge Model engine and calculates the accuracy of the Knowledge Model. Finally, the Prediction Manager takes the prediction results

set from the Knowledge Model engine and calculates the predictions of the Knowledge Model.

Col. 19, lines 4-52

3. The Study.

The "Study" is the basic unit the user creates to tackle a new data mining problem. In the disclosed embodiment, a Study may be created, opened, saved, saved as another Study, and modified. Study creation is an iterative process in which each new Discovery may give one new ideas to investigate, and thus new specifications to optimize one's Study. This could be as simple as selecting a new Output Parameter, or changing the categorization of a parameter. The Study use phase is the way to use the data mining tool when one considers the Knowledge Model as accurate. One can use it to discover new knowledge and/or to predict new situations. The processes underlying data are different if a Study is being built as compared to when an existing Study is being used, as shall be seen.

A Study is the container for Subject and Domain specifications, each Domain being a list of Data Sources, the Study allowing for setting Domains for each of the purposes of Discovery, Evaluation and/or Prediction, and describing the Parameters (variables or data fields), and indicating any specific Categorization. Thus, in this embodiment, a Study uses one and only one Subject, the Subject has Domains assigned for each of the tasks of Discovery, Evaluation and Discovery.

A Categorization represents the way a user partitions (abstracts) a numerical field in order to give a sense (semantic) to this field for the user's particular context. Depending on a user's needs, one might employ a more granular or more refined Categorization--it really depends on how one looks at the information. For example, a doctor might split the parameter "AGE" in several categories (because she knows that there are some diseases specific to each particular age): baby (ages 0-1), and toddler (ages 1-3). A toy company, on the other hand, might be interested only in these different categories: young children (age.ltoreq.4) and older children (ages 5-18). Thus, one may have the same data, yet categorize them differently depending on the purpose. One way to determine an appropriate Categorization is to analyze the data distribution, through histograms or otherwise. One may note clumping of data points in some areas, spreading out in others, this may suggest having partitions of unequal size, but roughly equal populations. This can be achieved by partitioning the data into very small sections, say 50 in number, so that the population in each section is not too great, with the population measured in each, and thus sections may be assembled together into larger "bins" as required to accomodate a useful Categorization. These bins need not be mutually exclusive, i.e., they may overlap.

Col. 20, lines 8-67

b. Describing the Domain.

Next, one describes the Domain. A Domain is a set of elements or data rows to which variables are restricted to be used by the data mining tool. In a Study, one may set one of several Domains: (i) the training or Discovery domain; (ii) the Evaluation domain; and (iii) the Prediction domain. Each Domain allows the user to realize a step in the study. As mentioned earlier, a Data Source can be a data file or a database request. It is not important to know exactly the nature of the data. The Data Source simply prompts the data mining tool as to how to obtain the data.

The training or Discovery domain is a historical data set with which the user wishes to base discovery. In a simple Study, under the disclosed embodiment, a training domain is a single database request or a flat file containing the data. For trend analysis study, the Discovery Domain could be composed of several requests or several data sets each representing a step of a trend in the historical data. By default, when the user defines one Discovery domain, the Evaluation Domain is set to the same set. The Prediction Domain may contain as many requests as the user desires, depending on the different predictions the user wishes to make.

In the "Domain" mode, one has the dialog box 2100b as shown in FIG. 10(b). The user thus selects the data source from the data source list 2116 and presses the Discovery button 2118 to set the Discovery Domain 2126. This data set will be run for the Knowledge Discovery Process to obtain the Discovery Results. Similarly, through use of the Evaluation button 2120 one sets the Evaluation Domain 2128, and the Prediction button 2122 will set the Prediction Domain 2130. Corrections to any of these sets may be made through the Remove button 2124.

Data Source Utilization refers to the range of data to be used from a Data Source. Thus, this may be specified through a lower and upper row range through data entry fields 2132 and 2134, respectively, or through a proportional cutoff as specified through data entry field 2136. In this embodiment, specification of row ranges 2132, 2134 will cause a corresponding readjustment in the display of the proportional cutoff 2136 and vice-versa.

c. Describing the Parameters.

In the "Parameters" mode, one has the dialog box 2100c as shown in FIG. 10(c). Buttons "Add" 2143, "Add All" 2145, "Ignore" 2147 and "Default" 2149 allow the user to switch and remove fields, or Parameters, between data windows 2140 and 2142. The Default button 2149 will set the Parameters in data window 2142 to those not identified as "Ignored." An Ignored Parameter is one deemed not to serve a useful purpose in the Discovery Process, as will later be described.

Pop-up menu 2144 is used to specify any Parameter as input or output. In the present embodiment, at most one output field may be specified as output. This is also called the "Objective" of the Study. For example, the Objective might be to qualify a bank "ACCOUNT" status (e.g., "Balanced," "Overdraft," "VISA Late Payment"). However, it is not necessary to specify an Objective in the disclosed embodiment, the system creating a "virtual" dimension in the Segmentation process described later.

Pop-up menu 2148 allows the user to specify whether a Parameter is discrete or continuous, and if discrete, the number of segments may be entered through data entry box 2146.

<u>Col. 9, line 41 through col. 12, line 54</u>

c. The Inference Propagation Mechanism.

The propagation mechanisms incorporated into various embodiments of neuroagents include: (i) forward propagation; (ii) backward propagation; (iii) spontaneous backward propagation; and (iv) "retropropagation of necessities." All these propagation mechanisms are asynchronous, meaning that the update of neuroagents are event-driven.

FIGS. 6(a)-(c) show the different inference propagation mechanisms within the same system. In FIG. 6(a), forward propagation, the usual mode of propagation,

is shown. Neuroagent 600 is connected to neuroagent 610 via the latter's Minimal Excitation Zone 614 or Contextual Excitation Zone 612 (connections 604 or 602, respectively). Thus, when neuroagent 600 is validated, this state is propagated, as signals 608 or 606, as the case may be, to neuroagent 610. Thus, if the minimal conditions on neuroagent 610 are satisfied and/or the excitation threshold reached, neuroagent 610 itself may be validated and the propagation may continue further.

FIG. 6(b) shows backward propagation. Notice that the connections are analogous to FIG. 6(a), with neuroagent 620 connected to neuroagent 630 via the latter's Minimal Excitation Zone 634 or Contextual Excitation Zone 632 (connections 624 or 622, respectively). Backward propagation may be performed as a result of an explicit selection to backward propagate, or may occur spontaneously, through the mechanism of Hypothesis. The Hypothesis mechanism triggers backward propagation where neuroagent validation is almost present, as for example: (i) the Minimal Excitation Zone 634 is validated and the Contextual Excitation Zone 632, is near but below its excitation threshold, say in the range of 80-100%; or (ii) Contextual

Excitation Zone 632 is validated but the Minimal Excitation Zone 634 is indeterminate. Thus, depending on the mechanism, neuroagent 630 will backward propagate, as either signals 636 or 638 (under Hypothesis, owing to the spontaneous generation by the Contextual 632 or Minimal 634 Excitation Zones, respectively), to neuroagent 620. Thus, neuroagent 620, based on this backward propagation, will find itself either validated, inhibited or indeterminate. The indeterminate state may cause further spontaneous backward propagation, or the process will stop if neuroagent 620 is not configured to go into Hypothesis.

Retropropagation of the necessities involves only the Minimal Excitation Zone, and is a means to verify implicit deductions, as shown in FIG. 6(c). Here, neuroagent 650 is connected to neuroagent 640 through the latter's Minimal Excitation Zone (connection 652). Neuroagent 650 may be connected to other neuroagents (not shown) through connection 654. Thus, if neuroagent 640 is validated (signal 642) retropropagation will occur (signal 644), thereby validating neuroagent 650, which will forward propagate itself (signal 656) as will neuroagent 640 (via connection 646). The implicit deductions are thus verified in the sense that the network connection topology supplies the information. Say that neuroagent 640 represents "CAR" and neuroagent 650 "WHEELS". Thus, this connection of neuroagents 640, 650 would supply the deduction that "CARS" implies "WHEELS".

3. Construction of Neuroagent Networks.

With the neuroagent approach, it is possible to design a knowledge base through either explicit modelling, learning, or both. This versatility enhances the quality of the knowledge bases, since in many cases neither explicit modelling nor learning from examples are sufficient of themselves.

The learning process is conducted with two objectives: (i) to automatically establish the connection weights as in usual connectionist models; and (ii) to automatically establish the topology of the network. Due to the neuroagent's connectionist architecture, the system will not be a "black box" at the end of the learning; rather, it will be able to reach semantic conclusions, i.e., make explicit predictions as to: minimal conditions for the validation of outputs, the simultaneous presence of certain inputs, and the specificity of certain inputs, etc.

> Say one had a medical database to study in order to design a knowledge base of pathologies diagnosed in various patients. At the end of the learning process, one would obtain the connection weights established by the system, but also: (i) which symptoms are minimal (necessarily present) in order to diagnose a given pathology; (ii) which symptoms are always found together; and (iii) which symptoms are specific to a given pathology.

At the indicated locations, Pham merely describes predictive modeling and data mining, but nothing describe at these locations relates to "the derived measure" being "invoked within an application template that is a sequence of segments, filters, measures and functions linked together in a workflow." Indeed, nothing at these locations in Pham discusses anything equivalent to such an application template. Instead, Pham merely describes knowledge models generally and neuoragents specifically, but neither concept reads on application templates that are used to save sequences of a workflow. Consequently, the Pham reference does not teach or suggest all the limitations of Applicant's independent claims.

Thus, Appellant submits that independent claims 1, 15, and 29 are allowable over Pham. Applicant's attorney submits that dependent claims 2-14, 16-28, and 30-42 are allowable over Pham in the same manner as the independent claims, because they are dependent on independent claims 1, 15, and 29, respectively, and thus contain all the limitations of the independent claims. In addition, dependent claims 2-14, 16-28, and 30-42 recite additional novel elements not shown by Pham .

IV. CONCLUSION

In view of the above, it is submitted that this application is now in good order for allowance and such allowance is respectfully solicited.

Should the Examiner believe minor matters still remain that can be resolved in a telephone interview, the Examiner is urged to call Applicant's undersigned attorney.

Respectfully submitted,

GATES & COOPER LLP
Attorneys for Applicants

Howard Hughes Center
6701 Center Drive West, Suite 1050
Los Angeles, California 90045
(310) 641-8797

Date: September 18, 2006

By: GHG

Name: George H. Gates
Reg. No.: 33,500

GHG/